Available online at www.sciencedirect.com
ScienceDirect
Computers and Electronics in Agriculture 54 (2006) 22–35
Computers and electronics in agriculture
www.elsevier.com/locate/compag

Systematic approach for using hyperspectral imaging data to develop multispectral imagining systems: Detection of feces on apples

Alan M. Lefcout [a,*], Moon S. Kim [a], Yud-Ren Chen [a], Sukwon Kang [b]

[a] *USDA Agricultural Research Service, Instrumentation and Sensing Laboratory, Henry A. Wallace Beltsville Agricultural Research Center, Building 303 Powder Mill Road, BARC East, Beltsville, MD 20705, United States*
[b] *National Agricultural Mechanization Research Institute, Rural Development Administration, Suwon, Republic of Korea*

Received 4 January 2006; received in revised form 28 June 2006; accepted 28 June 2006

Abstract

The large size of data sets generated using hyperspectral imaging techniques significantly increases both the capability and difficulty of designing detection and classification systems. Of particular interest is the confluence with increasing use of multispectral imaging in machine vision, particularly in the area of food safety inspection. The purpose of this study was to develop a robust method for selecting one or two wavelengths for multispectral detection systems using hyperspectral data. The actual performance of detection algorithms in terms of true positives and false positives was used as optimization criteria. Detection of fecal contamination on apples is an important health safety issue. Prior observations suggest reflectance or fluorescence imaging in the visible to near-infrared can be used to detect such contamination. For this study, 1:2, 1:20, and 1:200 dilutions of dairy feces were applied to 100 Golden and 100 Red Delicious apples. Apples were imaged using a hyperspectral system, and a uniform power transformation was used to reduce inter-apple intensity variability. Detection was accomplished by applying a binary threshold to transformed single wavelength images and images construct using ratios or differences of images at two different wavelengths. Optimization criteria allowed for a maximum of three false positives. For reflectance imaging, maximum detection rates for 1:20 dilution spots on Golden and Red Delicious apples images were 100% and 62.5% using $R816 - R697$ and $R784 - R738$, respectively. For fluorescence imaging, maximum detection rates for 1:200 dilution spots on Golden and Red Delicious apples were 97.9% and 58.3% using $F665/F602$ and $F647/F482$, respectively. In all case, more concentrated dilution spots were detected at 100%. Maximum detection rates for Red Delicious apples required use of a Prewitt edge-detection filter. In comparison, tests of wavelengths and algorithms identified in previous studies using statistical methods such as principal component analysis produced lower detection rates, mainly due to problems with false positives. The procedures used for developing detection algorithms are not specific to detecting feces on apples, and it is theoretically easy to extend the results to detection schemes involving many wavelengths. The problem is the classical dilemma of rapidly increasing computational time. Still, given the costs of thoroughly testing a candidate detection algorithm, the time maybe warranted. Furthermore, as machine vision systems are often limited to one or two wavelengths due to practical considerations including cost, exhaustive search algorithms based-on optimizing the output of candidate detection algorithms should be cost-effective.
Published by Elsevier B.V.

Keywords: Fluorescence; Reflectance; Hyperspectral imaging; Imaging; Fecal contamination

* Corresponding author. Tel.: +1 301 504 8450; fax: +1 301 504 9466.
E-mail address: alefcour@anri.barc.usda.gov (A.M. Lefcout).

0168-1699/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.compag.2006.06.002

1. Introduction

Hyperspectral imaging techniques allow collection of data sets that include spectral information for each spatial pixel location. Often, such hyperspectral data are used to identify a limited number of wavelengths to be used in practical, real-time, multispectral detection systems (Shaw and Manolakis, 2002; Cheng et al., 2004). In the area of food safety, a specific example is efforts to use hyperspectral reflectance or fluorescence data in the visible and near-infrared (NIR) regions to select appropriate wavelengths for detection of feces on apples (Kim et al., 2002a,b; Liu et al., in press) or cantaloupes (Vargas et al., 2005). Generally, selection of wavelengths was done by looking at characteristics of the data sets to select wavelengths and image transformations that enhanced the visible contrast in resulting images between contamination sites and uncontaminated surface areas. Only limited attempts have made to consider the actual percentage of contamination sites that could be detected, particularly in relation to false-positive rates (Lefcourt et al., 2003, 2005). The purpose of this study was to develop a robust method for selecting one or two wavelengths for a multispectral detection system using hyperspectral data, and to use actual detection performance to optimize the selection process. A secondary purpose was to directly compare the efficacy of using reflectance or fluorescence imaging for detection using the same samples imagined in exactly the same positions.

Detection of fecal contamination on apples is a critical health safety issue as feces can be the source of a number of human pathogens (Armstrong et al., 1996; Blackburn and McClure, 2002; Hui, 2001; Mead et al., 1999). The FDA has solicited development of systems to detect contaminated apples (FDA, 2001). The efficacy of algorithms to detect feces on apples has been examined using more restricted data sets generated using laser-induced fluorescence imaging (Lefcourt et al., 2003, 2005). These data sets were limited to images acquired at four wavelengths. Effective detection algorithms were found that utilized both single-band images and ratios of images at two wavelengths. Effectiveness was evaluated by considering the acceptable number of false positives. Contamination sites were detected using a binary threshold after the images had been suitably transformed. The best detection was obtained by subjecting images to a universal power transformation followed by edge detection (Lefcourt et al., 2005). The universal power transformation reduces normal apple-to-apple intensity variation and enhances the contrast between contamination sites and uncontaminated surface areas (Lefcourt and Kim, 2006). For this study, single-band images, and ratios or differences of two single-band images were examined. Two classes of detection schemes were proposed based-on use of the uniform power transformation with and without the addition of edge detection. As the proposed detection schemes are complex and non-linear, optimization required use of an exhaustive search routine. To test the validity of the optimization methods, detection algorithms were developed using one-half of the apples and then applied to the remaining apples.

2. Methods

Golden Delicious and Red Delicious apples were artificially contaminated with titrations of dairy feces using a standardized laboratory protocol and subsequently scanned using a hyperspectral imaging system for reflectance responses to white light and fluorescence responses to UV-A excitation. For each apple variety, one-half of the apples were used to identify single- and two-band detection algorithms. For validation, the most promising algorithms were applied to the remaining apples.

2.1. Apples and feces application

Apples were handpicked from crates of early or mid-season, tree-run, apples at the Rice Fruit Co. (Gardners, PA) and stored at 3 °C. Fresh cow feces were collected at the Beltsville Agricultural Research Center dairy, diluted 1:2, 1:20, and 1:200 by weight with de-ionized water, and a single 30 μl drop of each of the three dilutions was immediately applied to each apple (Fig. 1). Drops were applied using a pipette with disposable plastics tips; about 2 mm was cut-off each tip with a razor blade to allow passage of particulates. Apples were then returned to the storage refrigerator to allow applications to dry overnight. Dry matter content of undiluted feces was about 16%. Feces were applied to 100 apples of each variety. Only 96 apples for each variety were used for analyses, the remainder were used to substitute for problem apples, e.g. where the application ran.

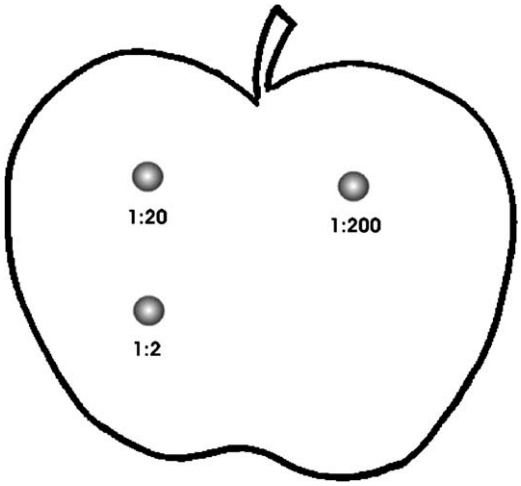

Fig. 1. Application sites for 30 μl of 1:2, 1:20, and 1:200 serial dilutions of dairy feces. The order of the applications followed a clockwise rotation based-on concentration. The initial quadrant for application of the 1:2 dilution was incremented one quadrant clockwise every fourth apple.

2.2. Hyperspectral imaging system

The imaging hardware consisted of a 16-bit camera (Model DV465, Andor Technology), a prism-grating system, and a zoom lens (2/3″; S6X11, Rainbow) mounted over a single-axis transition table (B-Slide, Velmex) along with UV-A and white light illumination sources (Fig. 2; Lefcourt and Kim, 2006). A hyperspectral data set is constructed from a series of images taken as an object is incrementally moved under the camera. The prism-grating system (ImSpector V9, Specim) has a spectral range from 430 to 900 nm, and uses a slit-width of 50 μm to produce a spectral resolution of about 4.5 nm. For UV-A illumination, two 30° light sources (Model XX-15A, Spectraline) are positioned 28 cm

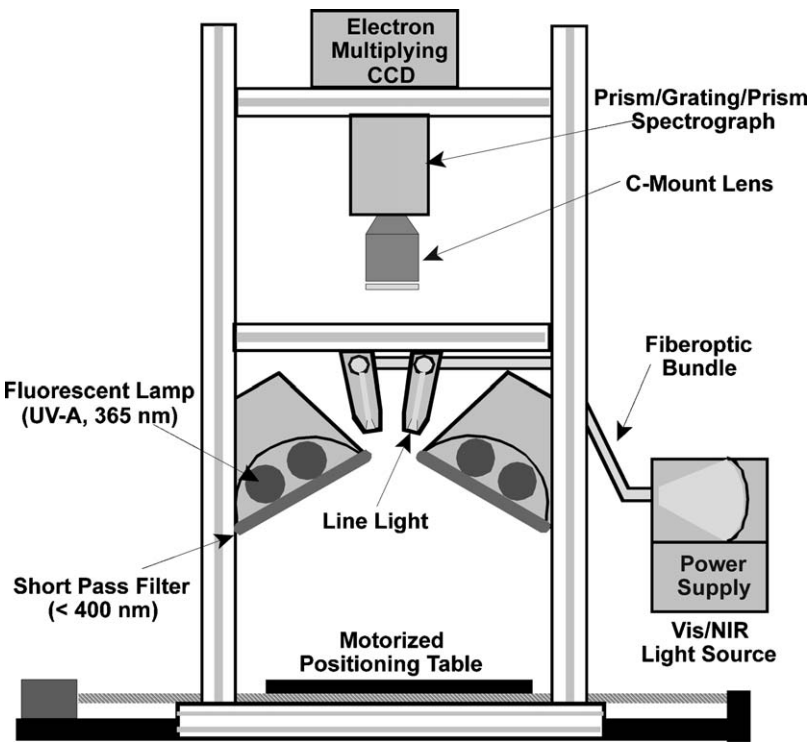

Fig. 2. Schematic representation of the hyperspectral imaging systems.

above the table and on each side of the imaging line at an angle of 18° relative to the vertical axis; output is restricted to between 300 and 400 nm using band-pass filters (UG-1, Schott Glass). Alternatively, white light is transmitted through two fiber optic bundles to randomly arranged rectilinear fiber bundles (Fiber-Lite A-240P, Dolan-Jenner Industries). Imaging parameters were set so that spatial resolution was about 0.5 mm/pixel and spectral resolution was about 4.4 nm/pixel. For fluorescence and reflectance imaging, data were examined from 452 to 729 nm and from 465 to 900 nm, respectively.

2.2.1. System operation

A program written in Visual Basic Version 6® (Microsoft) is used to control the transition table and acquire images, and for preliminary data analyses. After completion of an individual scan sequence, data were converted to hyperspectral image files using the ENVI® (Research Systems, Inc.) file format.

2.2.2. System calibration

The spectral response of the system was determined using known light sources as described by Kim et al. (2001). Tests indicated that it was not necessary to correct for individual pixel responsivity. However, a curvilinear response due to illumination was identified along the imaging axis. The curvilinear response resulted from the use of light sources that are the same dimension as the width of the translation table; hence, the outer extremes of the imaging area receive progressively less illumination. As the response was essentially flat in the middle two-thirds of the imaging axis, imaging for this study was limited to this area (∼24 cm). To meet this dimensional constraint, two parallel rows of six rubber cups were used to hold apples for imaging. Fluorescence images were corrected for black current and reflectance images were converted to percent reflectance as described previously (Kim et al., 2001).

2.3. Data analyses

The hyperspectral image files were analyzed using a series of programs written in Visual Basic Version 6® (Microsoft). All Basic programs relied on a library of functions developed in-house that allowed a range of image manipulations (e.g., brightness, contrast, normalization, exponential scaling, and functional combinations of images including ratios), filters (e.g., spatial, geometric, morphological, edge, and threshold; Weeks, 1996), and graphic displays such as false color images and histograms.

2.3.1. Masks to differentiate apple surfaces from background

To facilitate automated comparison of detection algorithms, masks to differentiate apple surfaces from background were created using selected images and binary thresholds. Minor inconsistencies in the apple masks, such as a void within the surface of an apple, were addressed using a pixel editor. Only one mask was created for each apple, and the same masks were used for all analyses.

2.3.2. Uniform power transformation

Gray-scale images for individual apples were adjusted for uniform power using a linear function to map pixel intensities based on the corresponding cumulative intensity histogram. For this study, the intensity corresponding to the 5% histogram level was mapped to an intensity of 20 and the 75% level to 65. As an individual apple occupied about 50% of the 187 by 187 pixel image area for each apple, the 75% intensity value corresponds to the median intensity of the surface of a generic apple. With a prism-based hyperspectral system, exposure time cannot easily be varied to account for differences in responsivity and actual intensity at different wavelengths. By using a camera with 16-bit resolution, it was possible to select a single exposure time that effectively resulted in a minimum resolution of 10 bits across wavelengths of interest. Transformed images were scaled to 10-bits by setting values less than 0 or greater than 1023 to 0 or 1023, respectively. This transformation reduces inter-apple intensity variability, differences due to effective sensitivity at different wavelengths, and it also improves the contrast between contaminated and uncontaminated apple surface areas (Lefcourt and Kim, 2006).

2.3.3. Creation of images for detection

Single-band images of individual apples used for all analyses were created by averaging three sequential wavelength images. Averaging was done to reduce noise and to approximate the throughput of 10 nm at half-max bandpass

Table 1
Procedures for image analysis

Before starting any analysis
Determine centroids of artificial contamination sites for each apple
Create a single-band apple image
Select apple image at the primary wavelength (Table 2)
Scale apple image using the uniform power transform
If enabled,
Apply edge-detection transform
Scale image using the uniform power transform
Create a two-band apple image
Select apple image at the primary wavelength (Table 2)
Scale apple image using the uniform power transform
Select apple image at the secondary wavelength (Table 2)
Scale apple image using the uniform power transform
Calculate two-band ratio or difference image
Scale two-band image using the uniform power transform
If enabled,
Apply edge-detection transform
Scale image using the uniform power transform

Procedures are used by analysis routine in Table 2.

filters (i.e., 3×4.4 nm = 13.1 nm). Ratio or difference images were created on a pixel-by-pixel basis using two different single-band images. Single-band, ratio, and difference images were scaled using the uniform transformation. A second set of images was created by subjecting the above images to a 5 by 5 Prewitt edge-detection filter. To reduce inappropriate detection of the edges of apples against the black background, pixel values less than 35 were set to 35 prior to edge detection. To keep binary thresholds consistent with those for the first set of images, output from the edge-detection filter was subject to the uniform power transformation. In addition, for single-band reflectance images where contamination sites were darker than uncontaminated apple surfaces, images were inverted and the background set to black prior to any transformation. It was not necessary to invert images used to construct ratio or difference images as all pair-wise permutations of single-band images were examined. A summary of procedures for creating images for analysis is in Table 1.

2.3.4. Locating contamination sites

A semi-automated software program was used to locate contamination sites on individual apples. Square masks (15 by 15 pixels) were placed so that each contamination site was centered within the appropriate mask. Different color masks were used for the three treatment levels. For each apple, a fourth mask was also placed in the untreated quadrant of the apple (Lefcourt and Kim, 2006).

2.3.5. Identification and detection of contamination and false-positive sites

Pixels in gray-scale images of individual apples with intensities greater than a selected binary threshold (varied for this study) and within a selected number of pixels of each other (five for this study) were considered to represent a single potential contamination site. For identification, the centroid of a potential contamination site was compared to the centroids of the previously created contamination masks for the apple. If the test centroid was within a selected pixel distance of a known centroid (five for this study), then the site was identified as the associated contamination site. If there was no match, the site was identified as a false positive. To allow visualization of this process, detected pixels could be color coded according to classification and superimposed on the original image. Information including the actual pixel counts for identified contamination sites and for the largest false-positive site for each apple was output for further analyses.

The final step for detection was to compare the pixel counts for identified and false-positives sites to a count threshold. Sites were considered to be detected if the associated pixel count exceeded the count threshold. For a given detection scheme, sorting the number of pixels for false-positive sites from smallest to largest allowed selection of the count threshold as a function of the number of false positives allowed. In this study, for automated comparisons of detection schemes, a maximum of three false positive was allowed, i.e., the count threshold was set to be the number

Table 2
Generalized algorithm for determining optimal detection parameters

```
Generalized analysis routine
  Set image type (single-band, ratio, or difference) and edge detection (yes or no)
    Increment primary wavelength
      Increment secondary wavelength if two-band image
        Increment apple number
          Create apple image (Table 1)
            Increment binary threshold
              Group adjacent pixels with intensities above the current binary threshold
              Determine centroid of each group
              Compare centroids of groups with known centroids of contamination sites
              Record pixel counts associated with contamination sites and with the largest false-positive group
            Binary threshold loop
        Apple number loop
        Sort records by pixel counts of largest false positive
        Determine count threshold that allows detection of a maximum of three false positives
        For each level of applied contamination, determine the number of apples where the pixel counts for identified contaminations sites
          exceed the detection threshold
        Record analysis parameters and number of detected contamination sites by concentration of applied feces
      Secondary wavelength loop for two-band images
    Primary wavelength loop
  Examine recorded data to determine the analysis parameters that resulted in the highest percentages of detected contamination sites
```

of pixels in the fourth-largest false-positive site. Thus, three false positives were allowed for each set of 48 apples subjected to a given detection scheme. In actuality, the number of false positives could be less than 3 if the pixel count for 4th to *n*th false positive was the same as the count for 1 of the first 3 false positives.

2.3.6. Automation of detection

To reduce the number of permutations of parameters that needed to be tested, a winnowing process was used whereby tests were initially conducted at intervals. Results of these tests were then used to select ranges of parameters for more detailed analysis. For the first pass, images were created at every third measured wavelength (at about 13.2 nm intervals) and analyzed with the binary threshold stepped from 60 to 140 at increments of 20. Results were summarized in terms of the percentage of contamination sites detected on apples for each level of treatment. For single-band images, three-dimensional graphs were created with wavelength on the *x*-axis, binary threshold on the *y*-axis, and percentage of sites detected on the *z*-axis. For ratio and difference image sets, the two wavelengths were on the *x*- and *y*-axes, and the maximum percentage of sites detected across binary thresholds was on the *z*-axis. For the second pass, the two or three most promising wavelengths, or pairs of wavelengths, for each apple variety were identified for single-band, ratio, and difference images both with and without edge detection. Irrespective of the type of analysis that produced the candidate waveform, all candidate waveforms for an apple variety were used to for second-pass tests, and the actual waveforms tested were expanded to include plus or minus five data points from the initial candidate waveforms. Binary thresholds were incremented at intervals of five from 15 below to 15 above values from pass 1 that produced optimal detection for the corresponding detection algorithm. The generalized routine for optimizing detection parameters is in Table 2.

2.3.7. Validation of detection parameters

Analyses of the remaining half of the apples were done using the most promising detection schemes. Results were compared to results for the first sets of apples in terms of percentages of contamination sites detected.

3. Results and discussion

Prior studies showed that the magnitude of fluorescent responses of feces peaked in the red region of the spectrum (Kim et al., 2002b), and that this region could be used to detect feces on apples (Kim et al., 2002b; Lefcourt et al., 2003). Similarly, reflectance responses in the NIR were demonstrated to be useful for detection of feces on apple (Kim et al.,

2002a; Liu et al., in press). In these studies, potential wavelengths for detection were selected using peaks and valleys in spectra of apples and of feces on apples, and using similar features in weighting functions for principal component images that visually appeared to differentiate feces-treated from untreated surface areas. In the current study, a range of potential detection algorithms were applied to images, and the best wavelengths were selected in terms of the actual ability to detect feces application sites on apples. The algorithms tested were ones that had previously been shown to enhance detection of fecal contamination sites on apples (Lefcourt et al., 2003, 2005; Lefcourt and Kim, 2006).

3.1. Problems for detection using reflectance imaging

3.1.1. Problems due to light source

The white light source was a linear array positioned above the apples. In the NIR, feces on apples appear darker than the surrounding surface. Unfortunately, the edges of the apple also appear dark due to the curvature of the apple. This phenomena prevented analysis of single-band images in terms of false positives as the edges of the apples were universally detected as false positives. Still, the 1:2 dilution sites were readily apparent at all NIR wavelengths. The 1:20 dilution sites were not apparent in most images. The problem of the edges was circumvented in ratio and difference images, where 100% of 1:2 dilution sites were detectable using appropriate wavelengths. The remaining question was whether ratios or differences could be used to detect 1:20 dilution sites. If so, the problem with edges could be considered to be moot as it is very unlikely that single-band images could be used to successfully detect 1:20 dilution sites.

3.1.2. Problems relating to calyx and stem

The stem and calyx regions have historically been major obstacles for using reflectance imaging to detect quality problems with apples; it has proved impossible to develop efficient algorithms to discriminate between problem sites and the stem or calyx (Bennedsen and Peterson, 2004; Kleynen et al., 2005; Penman, 2001; Throop et al., 2005). For reflectance imaging to be cost-effective, a practical method for orienting apples prior to imaging must be found so that these regions can be excluded from images. For this study, apples were oriented with the stem-calyx axis parallel to the imaging plane. Even so, these regions produced numerous false positives in ratio images. It is not clear whether this is an intrinsic problem, or whether a minor problem was exacerbated by lighting conditions as discussed above. Fortunately, the problem was substantially ameliorated in difference images. For this reason, only reflectance results for difference images are presented.

3.2. Problems for detection using fluorescence imaging

The major problem with using fluorescence to detect fecal contamination on apples is the difficulty of detecting sites where the feces visually obscures the underlying apple surface. It has been shown that there was no direct relation between the concentration of feces applied and the corresponding fluorescence response (Lefcourt et al., 2003, 2005). The measured responses for treated areas include contributions from the feces and from the underlying apple surface. Thick applications of feces, such as the 1:2 dilution, impede both the excitation and the response of the underlying surfaces, and high concentrations of fluorophors in the feces can reabsorb emissions from apple surfaces. In contrast, measured responses for thin applications, such as the 1:20 dilutions, which were visually translucent, are probably a summation of the responses of the feces and the underlying apple surfaces.

3.3. Intermediate results

Second-pass estimates of single-band detection rates for Golden Delicious apples using fluorescence images are shown in Fig. 3. As expected, the rates for detection of 1:2 dilution sites were much lower than rates for detection of less concentrated sites. The pattern and magnitude of detection rates were very similar when edge detection was added. These results clearly show that 1:2 dilution sites cannot be reliably detected using single-band fluorescence images. The second-pass estimates of detection rates using ratios of fluorescence images for Red Delicious apples are shown in Fig. 4. The detection algorithm included edge detection. Detection results for Red Delicious apples were improved through the use of edge detection for both reflectance and fluorescence images. Edge-detection algorithms look at the magnitude of localized changes in intensity; thus, reducing false positives due to natural variability and gradients of intensities measured across the surface of individual apples. In general, variability is greater for Red Delicious than

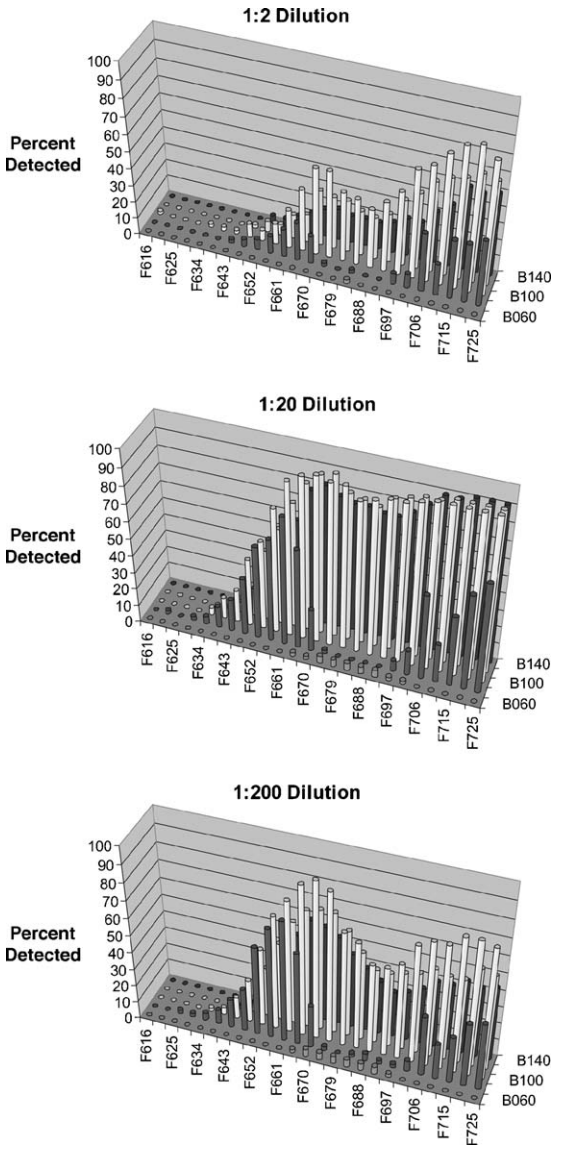

Fig. 3. Detection rates of contamination spots of dairy feces by dilution using single-band fluorescence images of Golden Delicious apples following application of the uniform power transform. Detection rates are shown as a function of wavelength (*F*) and binary threshold (*B*). Note difference in the ability to detect 1:2 and 1:20 dilution sites.

for Golden Delicious apples. For Golden Delicious apples, edge detection had only a small, positive, influence on detection. The increased detection rates occurred primarily when wavelengths adjacent to the optimal wavelengths were used for detection, and the increased computational cost of edge detection would probably not be warranted for practical implementation of a detection system. For these reasons, only results without edge detection are presented for Golden Delicious apples and with edge detection for Red Delicious apples. For fluorescence images, detection rates for difference images were very similar, but a little inferior, to those for ratio images. Thus, only results for ratio images are presented. However, the use of difference images should not be completely discounted, as differences are generally easier and faster to compute compared to ratios. It may be that a cost-effective solution to a specific problem could be based on using differences.

The iteration process identified a number of potential wavelength-pairs that allowed detection of fecal contamination spots. In some cases, the detection of 1:200 contamination spots was enhanced at the cost of detection of some of the 1:2 contamination spots. The criteria for selecting the optimal wavelength-pairs were extended to require that 100% of

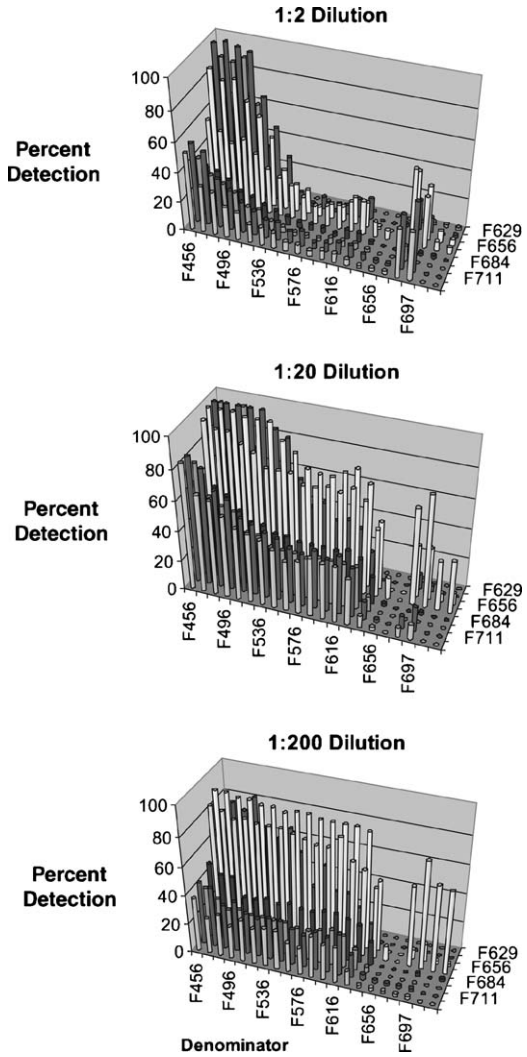

Fig. 4. Detection rates of contamination spots of dairy feces by dilution using ratios of fluorescence images of Red Delicious apples at selected wavelengths (F) following application of the Prewitt edge-detection filter.

the more concentrated dilution spots be detected. This extension allowed a single best wavelength-pair to be identified for each type of imaging and apple cultivar. It should be noted that, except in terms of detecting 1:2 dilution spots, there were no overlaps in "near" optimal wavelength-pairs that would allow detection using the same parameters for both apple cultivars.

3.4. Validation of selected wavelengths

Test and validation results for Golden Delicious apples using reflectance imaging are shown in Fig. 5. The $R816 - R697$ difference yielded the best detection results for 1:20 dilution sites: 93.3% and 100% for test and validation sets, respectively. The validation results were obtained using the some binary threshold and same count threshold that resulted from optimization using the test data. Detection for 1:2 sites was 100% in both cases. In subsequent discussions, it should be assumed that detection rates for more concentrated sites than the sites being discussed were 100%. The $R784 - R738$ difference proved best for using reflectance images to detect 1:20 dilution sites on Red Delicious apples (Fig. 6). Detection rates were 60.4% and 62.5% for test and validation data sets, respectively.

For fluorescence images, the ratio $F665/F602$ was best for detecting 1:200 dilution sites on Golden Delicious apples (Fig. 7). Detection rates were 97.9% for both test and validation datasets. For Red Delicious apples, detection rates for

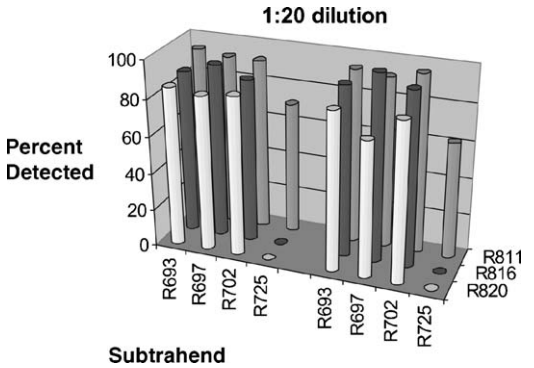

Fig. 5. Optimal reflectance wavelengths (R) for detecting 1:20 dilutions spots of dairy feces on Golden Delicious apples. The difference of $R816 - R697$ allowed detection of 93.7% of 1:20 dilution sites for the test data set (left) and 100% for the validation data set (right). Results are shown for images subject to the uniform power transformation. Detection rates for both test and validation data sets were 100% for 1:2 dilution sites.

the ratio $F665/F602$ were 72.9% and 58.3% for test and validation data sets, respectively (Fig. 8). For both reflectance and fluorescence images, the difficulty of detecting contamination sites for Red Delicious compared to Golden Delicious apples is reflected in the lower 1:200 detection rates; it also evidenced by the relatively lower, erratic, detection rates when wavelengths adjacent to the selected wavelengths are used for detection. In contrast, results for Golden Delicious apples were so robust that good detection results could be obtained using any of the adjacent wavelengths for detection.

In general, detection with reflectance imaging was about one-order of magnitude less sensitive than detection with fluorescence imaging. In addition, detection for Golden Delicious apples was better than detection for Red Delicious apples, regardless of imaging method, and detection for Red Delicious apples required use of edge detection. Despite lower sensitivity, reflectance imaging may prove to be more practical as reflectance imaging is less expensive to implement and can be used to simultaneously detect quality problems. In fact, it may prove to be impractical to attempt to differentiate between detection of fecal contamination and other quality problems.

3.5. Comparison of selected wavelengths with prior recommendations

For reflectance imaging, Liu et al. (in press) recommended using the ratio $R811/R725$ for detecting feces on apples. Figs. 5 and 6 include results for difference images constructed using these wavelengths and detection rates are clearly less than those for the selected wavelengths. Comparison of ratio results showed similar differences. Kim et al. (2002a) examined potential candidate wavelengths for different apple cultivars. However, none of the identified wavelengths correspond to the wavelengths selected in this study. For fluorescence images, Kim et al. (2003b) made a general recommendation for using a ratio of an image in the red band to an image in the blue or green band, and used the

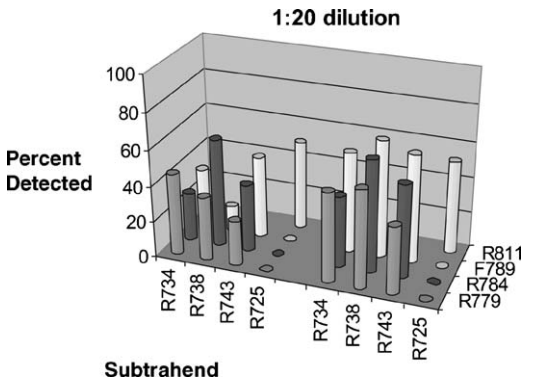

Fig. 6. Optimal reflectance wavelengths (R) for detecting 1:20 dilution spots of dairy feces on Red Delicious apples. The difference of $R784 - R738$ allowed detection of 60.4% of 1:20 contamination sites for the test data set (left) and 62.5% for the validation data set (right). Results are shown for images subject to Prewitt edge detection. Detection rates for both test and validation data sets were 100% for 1:2 dilution sites.

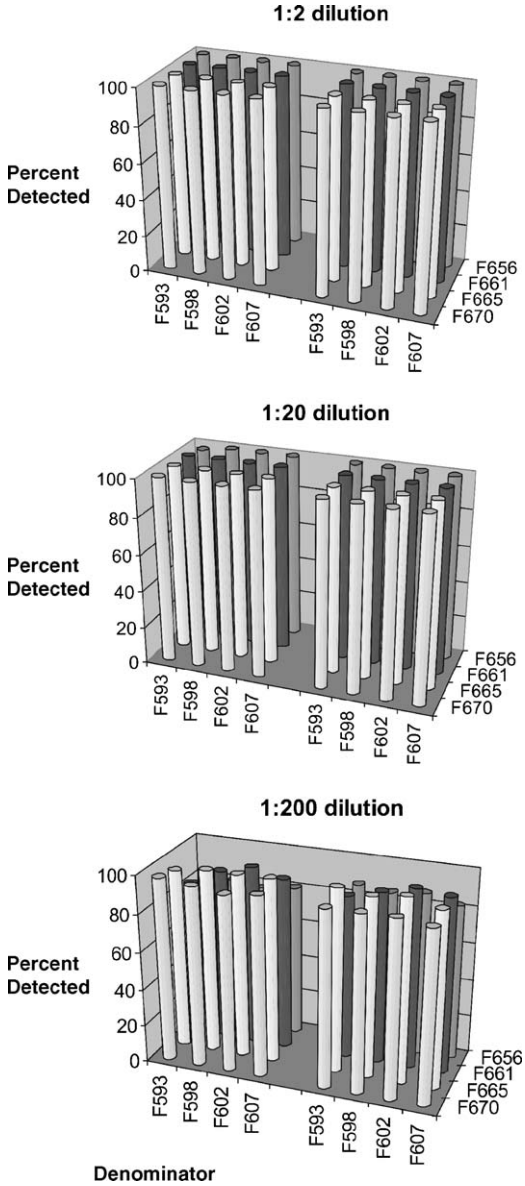

Fig. 7. Optimal fluorescence wavelength-pairs (F) for detecting spots of dairy feces by dilution on Golden Delicious apples. The ratio of $F665/F602$ allowed detection of 97.9% of 1:200 dilution sites for the test data set (left) and 97.9% for the validation data set (right). Results are shown for images subject to the uniform power transformation. Detection rates for both test and validation data sets were 100% for 1:2 and 1:20 dilution sites.

ratio of $F685/F450$ for demonstration as a promising wavelength-pair. While the recommendation is sound, the results of this study show that the selection of the appropriate wavelength for the denominator, as well as the numerator, is critical. The original supposition that selection of the exact wavelengths were not critical was based on the observation that the average spectra for apple and for feces on apples were smooth curves with no sharp irregularities.

The discrepancies between prior and current results are likely due to a number of factors. A primary factor is that the prior studies did not address the spatial relation of pixels with similar spectral characteristics. Also, a skewed probability distribution that favors detection will have a negative impact on statistical optimization methods that minimize some function of the variance. A third factor is the multi-modal nature of sites that cause false positives. False positives are often outliers, and selecting the best wavelengths based-on statistical measures can be dangerous unless the sources of outliers that result in false positives are well understood.

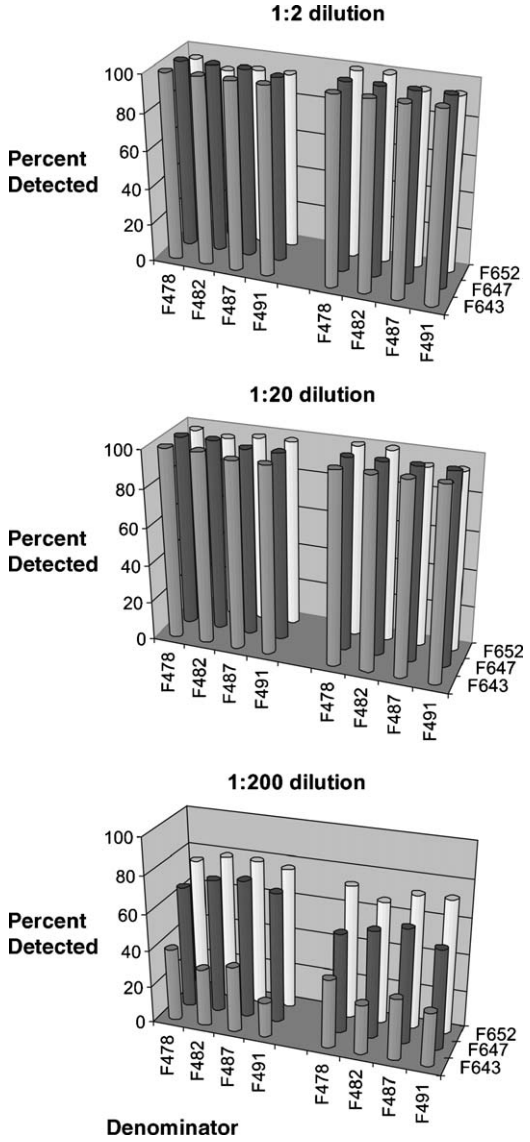

Fig. 8. Optimal fluorescence wavelength-pairs (*F*) for detecting spots of dairy feces by dilution on Red Delicious apples. The ratio of *F*647/*F*482 allowed detection of 72.9% of 1:200 dilution sites for the test data set (left) and 58.3% for the validation data set (right). Results are shown for images subject to the Prewitt edge-detection filter. Detection rates for both test and validation data sets were 100% for 1:2 and 1:20 dilution sites.

One of the reasons prior studies have emphasized ratio functions is to address differences in inter-apple intensity variability (Kim et al., 2002b). However, the uniform power transformation does a better job of reducing this variability (Lefcourt and Kim, 2006). As a corollary, pretreatment of images using the uniform power transformation prior to principal component analysis dramatically alters the principal component images and the associated weighting functions (personal observation). The results of this study suggest ratios, and differences, are important more for their ability to emphasize differences in spectral trends between treated and untreated sites.

3.6. Caveats

3.6.1. Possible errors due to instrumentation and source of apples

The results of this study are dependant to some extent on the instrumentation used and the sample set. The degree of this dependence is difficult to estimate. Differences in selected wavelengths for this and prior studies are unlikely to be

due to differences in instrumentation as all the studies are from this laboratory and utilized the same equipment. Still, replication of this experiment with different equipment might very well identify different optimal wavelength-pairs.

Apples for experimental trials vary by vendor and time of the year. Visual differences among sets of apples can be fairly dramatic, and it is easy to believe that these differences could impact the selection of the best wavelengths for detecting feces on the apples. Only multiple tests across many sets of apples can fully answer this question. In the worst case, it may be necessary to calibrate detection instrumentation for different sets of apples, or to design self-calibrating instruments that continuously respond to changes in apple characteristics.

3.6.2. Computational costs

The procedures used for developing detection algorithms are not specific to detecting feces on apples, and it is theoretically easy to extend the results to detection schemes involving many wavelengths. The problem is the classical dilemma of rapidly increasing computational time. Still, given the costs of thoroughly testing a candidate detection algorithm, the time maybe warranted. Furthermore, as machine vision systems are often limited to one or two wavelengths due to practical considerations including cost, exhaustive search algorithms based-on optimizing the output of candidate detection algorithms should be cost-effective.

4. Conclusion

In the current study, hyperspectral fluorescence and reflectance data sets were constructed for Golden and Red Delicious apples treated with 1:2, 1:20, and 1:200 dilutions of dairy feces. A of a range of potential detection methods were applied to single-band, ratio and difference images, and the best wavelengths and algorithms for detecting the contamination sites were selected in terms of the actually ability to actually detect feces on apples with a limited number of false positives. Selected wavelengths and associated algorithms that correctly identified 100% of 1:2 dilution sites using reflectance imaging and 100% of 1:2 and 1:20 dilution sites using fluorescence imaging were validated using an independent data set. The method developed for identifying appropriate wavelengths can be applied in many areas besides detecting feces on apples, and will be of interest to scientists and engineers faced with the task of reducing hyperspectral data to a more usable format.

References

Armstrong, G.L., Hollingsworth, J., Morris Jr., J.G., 1996. Emerging foodborne pathogens: *Escherichia coli O157:H7* as a model of entry of a new pathogen into the food supply of the developed world. Epidemiol. Rev. 18, 29–51.

Bennedsen, B.S., Peterson, D.L., 2004. Identification of apple stem and calyx using unsupervised feature extraction. Trans. ASAE 47 (3), 889–894.

Blackburn, C.W., McClure, P.J. (Eds.), 2002. Foodborne Pathogens: Hazards, Risk Analysis, and Control. CRC Press, Cambridge.

Cheng, X., Chen, Y.R., Yang, T., Wang, C.Y., Kim, M.S., Lefcourt, A.M., 2004. Hyperspectral imaging and feature extraction methods in fruit and vegetable defect inspection. Trans. ASAE 47 (4), 1313–1320.

Food Drug Administration (FDA), 2001. Hazard analysis and critical control point (HAACP); procedures for the safe and sanitary processing and importing of juices. Fed. Reg. 66, 6137–6202.

Hui, Y.H. (Ed.), 2001. Foodborne Disease Handbook. Marcel Dekker, New York.

Kim, M.S., Chen, Y.R., Mehl, P.M., 2001. Hyperspectral reflectance and fluorescence imaging system for food quality and safety. Trans. ASAE 44, 721–729.

Kim, M.S., Lefcourt, A.M., Chao, K., Chen, Y.R., Chan, D.E., Intaek, K., 2002a. Multispectral detection of fecal contamination on apples based on hyperspectral imagery. Part I. Application of visible-near infrared reflectance imaging. Trans. ASAE 45 (6), 2027–2037.

Kim, M.S., Lefcourt, A.M., Chen, Y.R., Kim, I., Chao, K., Chan, D., 2002b. Multispectral detection of fecal contamination on apples based on hyperspectral imagery. II. Application of fluorescence imaging. Trans. ASAE 45, 2027–2038.

Kleynen, O., Leemans, V., Destain, M.F., 2005. Development of a multi-spectral vision system for the detection of defects on apples. J. Food Eng. 69 (1), 41–49.

Lefcourt, A.M., Kim, M.S., Chen, Y.R., 2003. Automated detection of fecal contamination of apples by multispectral laser-induced fluorescence imaging. Appl. Opt. 42, 1–9.

Lefcourt, A.M., Kim, M.S., Chen, Y.R., 2005. Detection of fecal contamination in apple calyx by multispectral laser-induced fluorescence. Trans. ASAE 48, 1587–1593.

Lefcourt, A.M., Kim, M.S., 2006. Technique for normalizing intensity histograms of images when the approximate size of the target is known: Detection of feces on apples using fluorescence imaging. Comput. Electron. Agric. 50, 135–147.

Liu, Y., Chen, Y.R., Kim, M.S., Chan, D.E., Lefcourt, A.M., in press. Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. J. Food Eng.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., Tauxe, R.V., 1999. Food-related illness and death in the United States. Emerg. Infect. Dis. 5, 607–625.

Penman, D.W., 2001. Determination of stem and calyx location on apples using visual inspection. Comput. Electron. Agric. 33 (1), 7–18.

Shaw, G., Manolakis, D., 2002. Signal processing for hyperspectral image exploitation. IEEE Signal Process. Mag. 19 (1), 12–16.

Throop, J.A., Aneshansley, D.J., Anger, W.C., Peterson, D.L., 2005. Quality evaluation of apples based on surface defects: development of an automated inspection system. Postharvest Biol. Technol. 36 (3), 281–290.

Vargas, A.M., Kim, M.S., Tao, Y., Lefcourt, A.M., Chen, Y.R., Luo, Y., Song, Y., Buchanan, R., 2005. Detection of fecal contamination on cantaloupes using hyperspectral fluorescence imagery. J. Food Sci. 70 (8), E471–E476.

Weeks Jr., A.R., 1996. Fundamentals of Electronic Image Processing. SPIE Optical Engineering Press, Bellington, WA.